IDENTIFICATION OF PARAMETERS IN DISTRIBUTED PARAMETER SYSTEMS BY REGULARIZATION

Costas Kravaris and John H. Seinfeld

Department of Chemical Engineering
California Institute of Technology
Pasadena, California 91125

Abstract

Identification of spatially varying parameters in distributed parameter systems from noisy data is an ill-posed problem. The concept of regularization, widely used in solving linear Fredholm integral equations, is developed for the identification of parameters in distributed parameter systems. A general regularization identification theory is first presented and then applied to a parabolic identification problem. Methods for the numerical implementation of the regularization identification approach are also presented.

1. Introduction

Consider the following distributed parameter dynamic system:

$$\frac{\partial u}{\partial t} + A(t)u = f, \qquad \text{in } \Omega \times]0,T[$$

$$u(x,0) = u_0 \quad , \qquad \text{in } \Omega$$

$$B_j u = g_j \qquad , \ j = 0,\dots, m-1, \text{ on } \Gamma \times]0,T]$$

where $\Omega \subset \mathbb{R}^n$ with boundary Γ and $0 < T < \infty$ and where

$$A(t)\ u = \sum_{|p|,|q|\leq m} (-1)^{|p|} D_x^p (a_{pq}(x,t)\ D_x^q u)$$

$$B_j u = \sum_{|h|\leq m_j} b_{jh}(x,t) D_x^h u \quad , \ j = 0,\dots, m-1$$

with $0 \leq m_j$ = order of $B_j \leq 2m-1$.

The parameter identification problem associated with the above dynamic system can be stated as follows: Assuming the input function f, the initial condition and the boundary condition(s) to be known, and given an observation of u, determine the system operator A(t), i.e. the parameters $a_{pq}(x,t)$.

A number of important physical identification problems fall within the above framework. For example, the partial differential equation

$$\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x,y)\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(\alpha(x,y)\frac{\partial u}{\partial y}\right) = f(x,y,t)$$

governs the temperature distribution in an inhomogeneous solid or the pressure distribution in a fluid-containing porous medium. In the case of fluid flow in a porous medium, α is termed the transmissivity. For models of petroleum reservoirs and subsurface aquifers the transmissivity is generally inaccessible to direct measurement, and its value must be inferred from measurements of the pressure u at wells.

The key difficulty in developing successful numerical techniques for identifying spatially-dependent parameters in partial differential equations is the fact that such problems are ill-posed in the sense of Hadamard ([8]). It can be shown by homogenization theory [1] that differential operators with highly oscillatory coefficients can be "replaced" by very different operators and still yield practically the same response. Lions [9] has, in fact, cited the main difficulty in identifying distributed coefficients in partial differential equations as preventing excess of oscillations in the coefficients.

The idea of regularization of ill-posed problems was first proposed by Tikhonov ([13], [14]) as a method of solving linear Fredholm integral equations of the first kind. Further development of the theory for ill-posed linear operator equations followed ([12]). Modern practical numerical methods for the solution of linear Fredholm integral equations involve regularization ([16]).

The object of the present work is to develop a regularization theory for the identification of parameters in distributed parameter systems. In §2 an abstract identification theory is developed. In §3 the theory is applied to a parabolic identification problem. §4 suggests methods for the numerical implementation of the regularization identification approach.

2. Identification by Regularization-Abstract Theory

Let $\mathcal{A}$, U and F be Banach spaces. $\mathcal{A}$ represents a space of partial differential operators, U represents the space of solutions and F the space of right hand sides. Consider a system described by

$$\Psi(A,u) = f \qquad (2.1)$$

where Ψ is a mapping, not necessarily linear, from $\mathcal{A} \times U$ into F. We assume:

(A1) Ψ is of C^k-class ($k \geq 1$)

(A2) There is an open subset $\mathcal{A}_c$ of $\mathcal{A}$ and an open subset U_c of U such that $\forall A \in \mathcal{A}_c$ (2.1) admits a unique solution $u \in U_c$.

(A3) $\forall A \in \mathcal{A}_c \quad \forall u \in U_c \ \frac{\partial \Psi}{\partial u}(A,u)$ is a linear homeomorphism of U onto F.

Thus, one can define an implicit function $u = \Phi(A)$ as the solution of (2.1). Φ is of C^k-class from $\mathcal{A}_c$ into U_c. Its first derivative $\Phi'(A)$ associates $\delta A \in \mathcal{A} \rightarrow \delta u \equiv \Phi'(A) \cdot \delta A \in U$, where δu is the solution of

$$\frac{\partial \Psi}{\partial u}(A,u) \cdot \delta u + \frac{\partial \Psi}{\partial A}(A,u) \cdot \delta A = 0 \qquad (2.2)$$

Furthermore, consider that A depends on a set of parameters λ belonging to the Banach space Λ. The set of physically admissible λ is $\Lambda_{ad} \subseteq \Lambda$. We assume:

(A4) $A;\ \Lambda \to \mathscr{A}$ is of C^k-class $(k \geq 1)$

(A5) Λ_{ad} is a norm-closed convex subset of Λ

(A6) $A(\Lambda_{ad}) \subseteq \mathscr{A}_c$

Now the identification problem can be posed as follows:

Knowing the mappings $\Psi;\ \mathscr{A} \times U \to F$ and $A;\ \Lambda \to \mathscr{A}$ and the element $f \in F$ and given an observation of u, find $\lambda \in \Lambda_{ad}$ to satisfy (2.1).

We need to be precise about the nature of the observation of u. Thus, consider a Hilbert Space $\mathscr{H}$ (Observation Space). Denote by $\Lambda_{\mathscr{H}}$ the canonical isomorphism of $\mathscr{H}$ onto $\mathscr{H}'$. Also, consider an observation operator, not necessarily linear, $\mathscr{C};\ U \to \mathscr{H}$ and assume

(A7) $\mathscr{C}$ is of C^k-class $(k \geq 1)$

The situation is depicted in Figure 1.

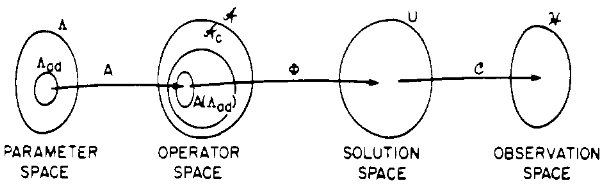

Figure 1

Thus, the identification problem can be viewed as solving in Λ_{ad} the (nonlinear) operator equation

$$(\mathscr{C} \circ \Phi \circ A)(\lambda) = z_d \tag{2.3}$$

If the operator $\mathscr{C} \circ \Phi \circ A;\ \Lambda_{ad} \to \mathscr{H}$ has a unique inverse and the inverse is continuous, one can, under certain conditions ([4]), apply the least-squares method ([3], [5]). It consists of minimizing over Λ_{ad} the functional

$$J_{LS}(\lambda) = \left\| \mathscr{C}\left(\Phi(A(\lambda))\right) - z_d \right\|^2_{\mathscr{H}} \tag{2.4}$$

As mentioned in the introduction, the identification of spatially-varying parameters in distributed parameter systems is, as a rule, an ill-posed problem (non-continuous dependence on the data). In other words, the problem of solving (2.3) is ill-posed. Hence, minima of $J_{LS}(\lambda)$ over Λ_{ad} (if any) will not depend continuously on the data z_d.

In order to regularize the parameter λ, we introduce a more regular space $\mathscr{R}$, for which we assume:

(A8) $\mathscr{R}$ is a Hilbert space.

(A9) $\mathscr{R}$ is densely imbedded in Λ.

(A10) The imbedding operator from $\mathscr{R}$ into Λ is compact.

Define $\mathscr{R}_{ad} = \mathscr{R} \cap \Lambda_{ad}$. With (A5) and (A9) is readily follows that $\mathscr{R}_{ad}$ is a norm-closed convex subset of $\mathscr{R}$.

We now introduce the *stabilizing functional*

$$J_S(\lambda) = \|\lambda\|^2_{\mathscr{R}}\ , \qquad \lambda \in \mathscr{R}_{ad} \tag{2.5}$$

and the *smoothing functional*

$$J_\beta(\lambda) = J_{LS}(\lambda) + \beta J_S(\lambda)\ , \quad \lambda \in \mathscr{R}_{ad} \tag{2.6}$$

where $\beta > 0$ is the *regularization parameter*. Identification by regularization proceeds as follows. Given $z_d \in \mathscr{H}$ and $\beta > 0$, find $\lambda_\beta \in \mathscr{R}_{ad}$ so as to minimize $J_\beta(\lambda)$.

THEOREM 2.1: *Assume that* (A1)-(A4) *and* (A6)-(A9) *are valid. Then the functional*

$$J_\beta(\lambda) = \left\|\mathscr{C}\left(\Phi(A(\lambda))\right) - z_d\right\|^2_{\mathscr{H}} + \beta\|\lambda\|^2_{\mathscr{R}} \tag{2.7}$$

is of C^k*-class. Its first derivative* $J'_\beta(\lambda);\ \mathscr{R} \to \mathbb{R}$ *is given by*

$$J'_\beta(\lambda)\cdot\delta\lambda = \left(\frac{\partial\Psi}{\partial A}(A(\lambda),u)\circ A'(\lambda)\cdot\delta\lambda, \phi\right)_{FF'}$$

$$+ 2\beta(\delta\lambda,\lambda)_{\mathscr{R}} \tag{2.8}$$

where u *is the solution of* $\Psi(A(\lambda),u) = f$ *and* ϕ *is the solution of*

$$\left[\frac{\partial\Psi}{\partial u}\left(A(\lambda),u\right)\right]^* \phi = -2[\mathscr{C}'(u)]^* \Lambda_{\mathscr{H}}\left(\mathscr{C}(u) - z_d\right) \tag{2.9}$$

Proof: That $J_\beta(\lambda)$ is of C^k-class is an immediate consequence of (A1)-(A4), (A7) and (A9).

Existence and uniqueness of the solution of $\Psi(A(\lambda),u) = f$ is guaranteed by (A2) and (A6).

Existence and uniqueness of the solution of (2.9) follows from the following facts:

(a) $-2[\mathscr{C}'(u)]^* \Lambda_{\mathscr{H}}(\mathscr{C}(u) - z_d) \in U'$,

(b) $\left[\frac{\partial\Psi}{\partial u}(A(\lambda),u)\right]^*$ is a linear homeomorphism of F' onto U', as a result of (A3).

The formula for the derivative of $J_\beta(\lambda)$ follows easily frrom (2.2) and (2.9).

THEOREM 2.2: *Under assumptions* (A1)-(A10), *the functional*

$$J_\beta(\lambda) = \left\|\mathscr{C}\left(\Phi(A(\lambda))\right) - z_d\right\|^2_{\mathscr{H}} + \beta\|\lambda\|^2_{\mathscr{R}}$$

admits a global minimum on $\mathscr{R}_{ad}$.

Proof: Let $\{\lambda_n\}$ be a minimizing sequence and $m = \inf_{\lambda \in \mathscr{R}_{ad}} J_\beta(\lambda)$. It is easy to see that $\{\lambda_n\}$ is norm-bounded in $\mathscr{R}$. Hence, there is a subsequence $\{\lambda_{n_k}\}$ that converges in the weak topology of $\mathscr{R}$ to some $\bar\lambda \in \mathscr{R}_{ad}$. By (A10) $\{\lambda_{n_k}\}$ converges to $\bar\lambda$ in the strong topology of Λ as well. Finally, using the continuity of the functional $J_{LS}(\lambda)$ in the norm-topology of Λ and the weak lower semicontinuity of $J_s(\lambda)$ in $\mathscr{R}$, it is not difficult to see that $m = J_\beta(\bar\lambda)$.

PROPOSITION 2.1: *A necessary condition for* $\lambda \in \mathscr{R}_{ad}$ *to be global minimum of* $J_\beta(\lambda)$ *on the set* $\mathscr{R}_{ad}$ *is*

$$J'_\beta(\lambda)\cdot(\nu-\lambda) \geq 0 \quad \forall \nu \in \mathscr{R}_{ad}$$

So far we have established existence of a minimum of the smoothing functional on $\mathscr{R}_{ad}$ and have given a necessary condition for optimality. Now we will show that minima of J_β depend continuously on the observation. This will be the key result of the regularization approach. Roughly speaking, what the next theorm says is the following:

Let $\tilde{\lambda}$ be the "true" value of the parameter and $\tilde{z}_d = \mathscr{C}\left(\Phi(A(\tilde{\lambda}))\right)$, what we would have observed with a zero-error observation. Provided that

(i) $\tilde{\lambda}$ is the unique preimage of z_d

(ii) β is an appropriately chosen function of the observation error, any minimum of $J_\beta(\lambda)$ converges (in the norm of Λ) to $\tilde{\lambda}$, as the observation error tends (in the nrom of $\mathscr{H}$) to zero.

THEOREM 2.3: *For any* $\beta > 0$ *and* $z_d \in \mathscr{H}$, *denote by* $\lambda_\beta \in \mathscr{R}_{ad}$ *any minimum of* $J_\beta(\lambda)$ *on* $\mathscr{R}_{ad}$. *Also, denote by* T_{δ_1} *the class of functions that are nonnegative, nondecreasing and continuous on the interval* $[0,\delta_1]$. *Suppose*

$$\begin{cases} \tilde{z}_d \in \mathscr{H} \\ \exists \text{ a unique } \tilde{\lambda} \in \mathscr{R}_{ad} \text{ with } \tilde{z}_d = \mathscr{C}\left(\Phi(A(\tilde{\lambda}))\right) \end{cases}$$

Then $\forall \varepsilon > 0 \quad \forall B_1, B_2 \in T_{\delta_1}$ *with*

$$\left\{ \begin{matrix} B_2(0) = 0 \\ \dfrac{\delta^2}{B_1(\delta)} \leq B_2(\delta) \end{matrix} \right\}$$

$\exists \delta_0(\varepsilon,B_1,B_2) \leq \delta_1$ *such that* $\forall z_d \in \mathscr{H} \quad \forall \delta \leq \delta_0$

$$\|z_d - \tilde{z}_d\|_{\mathscr{H}} \leq \delta \Longrightarrow \|\lambda_\beta - \tilde{\lambda}\|_\Lambda \leq \varepsilon$$

for all β *satisfying* $\dfrac{\delta^2}{B_1(\delta)} \leq \beta \leq B_2(\delta)$.

Proof: From $\|z_d - \tilde{z}_d\|_{\mathscr{H}} \leq \delta$ it easily follows that

$$J_\beta(\lambda_\beta) \leq \delta^2 + \beta\|\tilde{\lambda}\|^2_{\mathscr{R}} \tag{2.10}$$

Hence

$$\|\lambda_\beta\|^2_{\mathscr{R}} \leq \frac{\delta^2}{\beta} + \|\tilde{\lambda}\|^2_{\mathscr{R}}$$

$$\leq \beta[B_1(\delta_1) + \|\tilde{\lambda}\|^2_{\mathscr{R}}]$$

So λ_β and $\tilde{\lambda}$ belong to the set

$$\hat{\Lambda} = \{\lambda \in \mathscr{R}_{ad} / \|\lambda\|_{\mathscr{R}} \leq [B_1(\delta_1) + \|\tilde{\lambda}\|^2_{\mathscr{R}}]^{\frac{1}{2}}\}$$

which is precompact in the norm-topology of Λ. It follows that

$$\forall \varepsilon > 0 \ \exists \gamma(\varepsilon) > 0 \text{ such that } \forall \hat{\lambda} \in \hat{\Lambda}$$

$$\|\mathscr{C}\left(\Phi(A(\hat{\lambda}))\right) - \tilde{z}_d\|_{\mathscr{H}} \leq \gamma \Longrightarrow \|\hat{\lambda} - \tilde{\lambda}\|_\Lambda \leq \varepsilon$$

Now (2.10) implies

$$\|\mathscr{C}\left(\Phi(A(\lambda_\beta))\right) - \tilde{z}_d\|^2_{\mathscr{H}} \leq \delta^2 + B_2(\delta)\|\tilde{\lambda}\|^2_{\mathscr{R}}$$

so that

$$\|\mathscr{C}\left(\Phi(A(\lambda_\beta))\right) - \tilde{z}_d\|_{\mathscr{H}} \leq \|\mathscr{C}\left(\Phi(A(\lambda_\beta))\right) - z_d\|_{\mathscr{H}} + \|z_d - \tilde{z}_d\|_{\mathscr{H}}$$
$$\leq \left(\delta^2 + B_2(\delta)\|\tilde{\lambda}\|^2_{\mathscr{R}}\right)^{\frac{1}{2}} + \delta$$

The function $\psi(\delta) = (\delta^2 + B_2(\delta)\|\tilde{\lambda}\|^2_{\mathscr{R}})^{\frac{1}{2}} + \delta$ is a continuous monotonically increasing function satisfying $\psi(0) = 0$. Hence, one can choose $\delta_0 = \psi^{-1}(\gamma(\varepsilon))$ and have $\|\mathscr{C}(\Phi(A(\lambda_\beta))) - \tilde{z}_d\|_{\mathscr{H}} \leq \gamma(\varepsilon) \quad \forall \delta \leq \delta_0$. Thus we see that for all β satisfying $\dfrac{\delta^2}{B_1(\delta)} \leq \beta \leq B_2(\delta)$, the inequality $\|z_d - \tilde{z}_d\|_{\mathscr{H}} \leq \delta$ implies the inequality $\|\lambda_\beta - \tilde{\lambda}\|_\Lambda \leq \varepsilon$.

It remains to indicate how to select the regularization parameter as a function of an upper bound δ on the observation error (i.e. $\|z_d - \tilde{z}_d\|_{\mathscr{H}} \leq \delta$)

Method 1. When an a priori upper bound on $\|\tilde{\lambda}\|_{\mathscr{R}}$ is known, i.e. $\|\tilde{\lambda}\|_{\mathscr{R}} \leq \Delta$, one can choose $\beta(\delta) = (\delta/\Delta)^2$. (When $\mathscr{R}$ is a Sobolev space, $\| \ \|_{\mathscr{R}}$ is a measure of smoothness.) We note first that this choice of β satisfies the assumptions of Theorem 2.3. Furthermore, if $\lambda_{\beta(\delta)}$ is a minimizer of

$$J_\beta(\lambda) = \|\mathscr{C}\left(\Phi(A(\lambda))\right) - z_d\|^2_{\mathscr{H}} + \left(\frac{\delta}{\Delta}\right)^2 \|\lambda\|^2_{\mathscr{R}}$$

on $\mathscr{R}_{ad}$, then

$$J_\beta(\lambda_{\beta(\delta)}) \leq 2\delta^2$$

hence

$$\|\mathscr{C}\left(\Phi(A(\lambda_{\beta(\delta)}))\right) - z_d\|_{\mathscr{H}} \leq \sqrt{2}\,\delta$$

$$\|\lambda_{\beta(\delta)}\|_{\mathscr{R}} \leq \sqrt{2}\,\Delta$$

i.e. regularized solutions satisfy the constraints up to a factor of $\sqrt{2}$.

Method 2: This method has been suggested by Tikhonov and Arsenin [15]. Their suggestion is to choose $\beta(\delta)$ so that

$$\|\mathscr{C}\left(\Phi(A(\lambda_{\beta(\delta)}))\right) - z_d\|_{\mathscr{H}} = \delta$$

where $\lambda_{\beta(\delta)}$ minimizes

$$J_\beta(\lambda) = \|\mathscr{C}\left(\Phi(A(\lambda))\right) - z_d\|^2_{\mathscr{H}} + \beta(\delta)\,\|\lambda\|^2_{\mathscr{R}}$$

Note that such a $\beta(\delta)$ in general need not exist. In [7] necessary and sufficient conditions for existence are given. Also note that such a selection of β does not verify the assumptions of Theorem 2.3. However, following the idea of its proof, one can show

THEOREM 2.4: *Suppose*

$$\begin{cases} \tilde{z}_d \in \mathscr{H} \\ \exists \text{ a unique } \tilde{\lambda} \in \mathscr{R}_{ad} \text{ with } \tilde{z}_d = \mathscr{C}(\Phi(A(\tilde{\lambda}))). \end{cases}$$

Then $\forall \varepsilon > 0 \quad \exists \delta_0(\varepsilon) > 0$ *such that* $\forall z_d \in \mathscr{H} \quad \forall \delta \leq \delta_0$

$$\|z_d - \tilde{z}_d\|_{\mathscr{H}} \leq \delta \Longrightarrow \|\lambda_{\beta(\delta)} - \tilde{\lambda}\|_\Lambda \leq \varepsilon$$

where $\begin{cases} \beta(\delta) \text{ denotes a regularization parameter} \\ \lambda_\beta(\delta) \text{ denotes a minimizer of } J_{\beta(\delta)}(\lambda) \text{ on } \mathscr{R}_{ad} \end{cases}$

satisfying $\|\mathscr{C}(\Phi(A(\lambda_{\beta(\delta)}))) - z_d\|_{\mathscr{H}} = \delta$

3. Application to a Parabolic Identification Problem

(a) Distributed Observation

Let

Ω a bounded open subset of $\mathbb{R}^n$

Γ the boundary of Ω, a C^1-manifold with Ω locally on one side of Γ

T a real number with $0 < T < \infty$

$Q = \Omega \times]0,T[$

$\Sigma = \Gamma \times]0,T[$

and consider the following parabolic system

$$\begin{aligned}
&\frac{\partial u}{\partial t} - \sum_{j,k=1}^{n} \frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial u}{\partial x_k}\right) + a_0(x)u = f(x,t), && \text{in } Q\\
&u(x,0) = u_0(x) && \text{, in } \Omega \qquad (3.1)\\
&\frac{\partial u}{\partial \nu} = 0 && \text{, on } \Sigma
\end{aligned}$$

The variational formulation of the above Neumann problem is as follows

$$\int_\Omega \frac{\partial u}{\partial t} v + \int_\Omega \sum_{j,k=1}^{n} a_{jk}(x)\frac{\partial u}{\partial x_k}\frac{\partial v}{\partial x_j} + \int_\Omega a_0(x)uv = \int_\Omega fv \qquad \forall v \in V$$

$$u(x,0) = u_0(x)$$

where $V = H^1(\Omega)$

We assume that the parameters a_{jk}, $a_0 \in C^0(\bar{\Omega})$, i.e. the parameter space

$$\Lambda = \left(\sum_{j,k=1}^{n} C^0(\bar{\Omega})\right) \times C^0(\bar{\Omega})$$

which is a Banach space with norm

$$\|\lambda\|_\Lambda = \max\left\{\|a_{jk}\|_{C^0(\bar{\Omega})}, \|a_0\|_{C^0(\bar{\Omega})}\right\}$$

The set of admissible parameters is taken to be

$$\Lambda_{ad} = \left\{\lambda \in \Lambda \,/\, \sum_{j,k=1}^{n} a_{jk}(x)\xi_j\xi_k \geq \kappa(\xi_1^2 + \ldots + \xi_n^2)\right.$$
$$\left.\forall \xi \in \mathbb{R}^n \;\; \forall x \in \Omega \text{ and } a_0(x) > \kappa_0 \quad \forall x \in \Omega\right\}$$

where κ and κ_0 are given positive numbers; so (A5) is satisfied.

Now given $\lambda \in \Lambda$ define $A \in \mathscr{A} = \mathscr{L}(V,V')$ by

$$(Au,v)_{V'V} = \int_\Omega \sum_{j,k=1}^{n} a_{jk}(x)\frac{\partial u}{\partial x_k}\frac{\partial v}{\partial x_j} + \int_\Omega a_0(x)uv \qquad \forall u,v \in V$$

The open subset of $\mathscr{A}$

$$\mathscr{A}_c = \left\{A \in \mathscr{A} \,/\, \exists \zeta > 0: \;(Av,v)_{V'V} \geq \zeta\|v\|_V^2 \quad \forall v \in V\right\}$$

is the set of coercive operators. It is straightforward to verify assumptions (A4) (with $k = \infty$) and (A6).

System (3.1) (variational formulation) can now be rewritten in a more compact form as follows:

$$\begin{aligned}
&\frac{du}{dt} + Au = f\\
&u(0) = u_0
\end{aligned} \qquad (3.2)$$

where we have used the notation $\frac{du}{dt}(v)$ in place of $\int_\Omega \frac{\partial u}{\partial t} v$ and $f(v)$ in place of $\int fv$. It is known ([10, p. 102]) that for every $A \in \mathscr{A}_c$, $f \in L^2(0,T;V')$ and $u_0 \in L^2(\Omega)$, (3.2) admits a unique solution $u \in U = W(0,T) = \{u/u \in L^2(0,T;V), \frac{du}{dt} \in L^2(0,T;V')\}$ which depends continuously on f and u_0.

Taking $F = L^2(0,T;V') \times L^2(\Omega)$, $U_c = U = W(0,T)$ and defining the mapping

$$\Psi;\; (A,u) \in \mathscr{A} \times U \to \left(\frac{du}{dt} + Au,\, u(0)\right) \in F$$

it is not difficult to verify assumptions (A1) (with $k = \infty$), (A2) and (A3).

Finally, suppose one wants to identify a_{jk} and a_0 from distributed observation i.e. by observing $u(x,t)$ in Q. Take

$\mathscr{H} = L^2(Q)$

$\Lambda_\mathscr{H}$ = identity

$\mathscr{C}$ = injection of $W(0,T)$ into $L^2(Q)$

$$\mathscr{R} = \left(\bigtimes_{j,k=1}^{n} H^\ell(\Omega)\right) \times H^\ell(\Omega) \text{ with } \ell > \frac{n}{2}$$

Thus (A7)-(A10) are automatically satisfied.

So the theory of section 2 is applicable and the results of Theorems 2.1-2.4 hold. Furthermore, from Theorem 2.1 one can deduce the following specific result:

THEOREM 3.1: *Given $z_d \in L^2(Q)$ and $\beta > 0$, the smoothing functional*

$$J_\beta(\lambda) = \int_Q [u(x,t;\lambda) - z_d(x,t)]^2\, dxdt + \beta\|\lambda\|_\mathscr{R}^2$$

where $u(x,t;\lambda) \in W(0,T)$ is the weak solution of (3.1), is of C^∞-class. Its first derivative is given by

$$J'_\beta(\lambda)\cdot\delta\lambda = \int_Q\left[\sum_{j,k=1}^{n}\delta a_{jk}\frac{\partial u}{\partial x_k}\frac{\partial p}{\partial x_j} + \delta a_0 up\right]dxdt + 2\beta(\delta\lambda,\lambda)_\mathscr{R}$$

where $p \in W(0,T)$ is the weak solution of the adjoint equation

$$\begin{cases}
\dfrac{\partial p}{\partial t} + \displaystyle\sum_{j,k=1}^{n}\frac{\partial}{\partial x_j}\left(a_{kj}(x)\frac{\partial p}{\partial x_k}\right) - a_0(x)p = 2(u(x,t) - z_d(x,t)), & \text{in } Q\\
\dfrac{\partial p}{\partial \nu} = 0 & \text{, on } \Sigma\\
p(x,T) = 0 & \text{, in } \Omega
\end{cases}$$

(b) Point Observation

Given a set of discrete points $x_1,\ldots,x_\mu \in \Omega$, we now consider the identification of the system (3.1) by observing $u(x_i,t)$, $i = 1,\ldots,\mu$. We have seen previously that the weak solution of (3.1) lies in $L^2(0,T;H^1(\Omega))$. Thus, for a weak solution u, the point value $u(x_i,t)$ has meaning if $H^1(\Omega) \subset C^0(\Omega) \Leftrightarrow n < 2$.

Since such an assumption is overly restrictive, we will consider here strong solutions, which lie in $H^{2,1}(Q)$ (hence they lie in $C^0(\bar{Q})$ for $n \leq 3$).

We will make strong regularity assumptions, such as

Γ is an (n-1)-dimensional C^2-manifold, with Ω locally on one side of Γ and $a_{jk} \in C^1(\bar{\Omega})$, $a_o \in C^0(\bar{\Omega})$, i.e. the parameter space

$$\Lambda = \left(\bigtimes_{j,k=1}^{n} C^1(\bar{\Omega}) \right) \times C^0(\bar{\Omega})$$

which is a Banach space with norm

$$\|\lambda\|_\Lambda = \max \left\{ \|a_{jk}\|_{C^1(\bar{\Omega})}, \|a_o\|_{C^0(\bar{\Omega})} \right\}$$

The set of admissible parameters is taken to be

$$\Lambda_{ad} = \left\{ \lambda \in \Lambda \Big/ \sum_{j,k=1}^{n} a_{jk}(x)\xi_j\xi_k \geq \kappa(\xi_1^2 + \ldots + \xi_n^2) \right.$$

$$\left. \forall \xi \in \mathbb{R}^n \quad \forall x \in \Omega \quad \text{and } a_o(x) \geq \kappa_o \quad \forall x \in \Omega \right\}$$

where κ and κ_o are given positive constants; so (A5) is satisfied.

As operator space we take

$$\mathscr{A} = \left\{ A \in \mathscr{L}\left(H^{2,1}(Q), L^2(Q)\right) / \right.$$
$$A = -\sum_{j,k=1}^{n} \frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial}{\partial x_k}\right) + a_o(x)$$
$$\left. \text{with } a_{jk} \in C^1(\bar{\Omega}) \text{ and } a_o \in C^0(\bar{\Omega}) \right\}$$

which is a Banach space with norm

$$\|A\|_{\mathscr{A}} = \max \left\{ \|a_{jk}\|_{C^1(\bar{\Omega})}, \|a_c\|_{C^0(\bar{\Omega})} \right\}$$

and denote by $\mathscr{A}_c$ its open subset

$$\mathscr{A}_c = \left\{ A \in \mathscr{A} / \{A - e^{i\theta}\frac{\partial^2}{\partial y^2}, \frac{\partial}{\partial \nu}\} \text{ is a regular elliptic system on } \bar{\Omega} \times \mathbb{R}_y \quad \forall \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \right\}$$

It is straightforward to verify assumptions (A4) (with $k = \infty$) and (A6).

It is known[1] ([11, p. 33]) that for every $A \in \mathscr{A}_c$, $f \in L^2(Q)$, $u_o \in H^1(\Omega)$ and $g \in H^{1/2,1/4}(\Sigma)$, the boundary-value problem

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} + Au = f && , \text{ in } Q \\ &u(x,0) = u_o && , \text{ in } \Omega \\ &\frac{\partial u}{\partial \nu} = g && , \text{ on } \Sigma \end{aligned}\right\}$$

admits a unique solution $u \in U = H^{2,1}(Q)$ that depends continuously on f, u_o and g.

Taking $F = L^2(Q) \times H^1(\Omega) \times H^{1/2,1/4}(\Sigma)$, and defining the mapping

$$\Psi; (A,u) \in \mathscr{A} \times U \to \left(\frac{\partial u}{\partial t} + Au, u(x,0), \frac{\partial u}{\partial \nu}\right) \in F$$

it is not difficult to verify assumptions (A1) (with $k = \infty$), (A2) and (A3).

Now to identify (a_{jk}) and a_o in (3.1) from an observation of u at the points x_i, $i = 1,\ldots,\mu$, take

$$\mathscr{H} = (L^2(0,T))^\mu$$

$$\Lambda_{\mathscr{H}} = \text{identity}$$

$$\mathscr{C};\ u(x,t) \in H^{2,1}(Q) \to (u(x_i,t),\ i = 1,\ldots,\mu) \in (L^2(0,T))^\mu$$

$$\mathscr{R} = \left(\bigtimes_{j,k=1}^{n} H^{\ell}(\Omega) \right) \times H^{\ell_o}(\Omega) \text{ with } \ell > 1 + \frac{n}{2}, \ \ell_o > \frac{n}{2}$$

Thus (A7)-(A10) are automatically satisfied.

So the theory of section 2 is applicable and the results of Theorems 2.1-2.4 hold. Furthermore, from Theorem 2.1 one can deduce the following specific result:

THEOREM 3.2: *Given* $z_d = (z_{d_1}(t),\ldots,z_{d_\mu}(t)) \in (L^2(0,t))^\mu$ *and* $\beta > 0$, *the smoothing functional*

$$J_\beta(\lambda) = \sum_{i=1}^{\mu} \int_0^T [u(x_i,t;\lambda) - z_{d_i}(t)]^2\, dt + \beta\|\lambda\|^2_{\mathscr{R}}$$

where $u(x,t;\lambda) \in H^{2,1}(Q)$ *is the strong solution of (3.1), is of* C^∞*-class. Its first derivative is given by*

$$J'_\beta(\lambda)\cdot\delta\lambda = \int_Q \left[- \sum_{j,k=1}^{n} \frac{\partial}{\partial x_j}\left(\delta a_{jk}\frac{\partial u}{\partial x_k}\right) + \delta a_o u\right] p\,dxdt + 2\beta(\delta\lambda,\lambda)_{\mathscr{R}}$$

where $p \in L^2(Q)$ *is the unique solution of*

$$\int_Q \left[\frac{\partial v}{\partial t} - \sum_{j,k=1}^{n} \frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial v}{\partial x_k}\right) + a_o(x)v\right] p\, dxdt$$

$$= -2\sum_{i=1}^{\mu}\int_0^T [u(x_i,t) - z_{d_i}(t)]v(x_i,t)dt$$

$$\forall v \in H^{2,1}(Q) \text{ satisfying } \frac{\partial v}{\partial \nu} = 0 \text{ on } \Sigma,\ v(x,0) = 0 \text{ in } \Omega$$

In other words, p is a distributional solution of

$$\frac{\partial p}{\partial t} + \sum_{j,k=1}^{n} \frac{\partial}{\partial x_j}\left(a_{kj}(x)\frac{\partial p}{\partial x_k}\right) - a_o(x)p$$
$$= 2\sum_{i=1}^{\mu}(u(x_i,t) - z_{d_i}(t)) \otimes \delta(x-x_i), \text{ in } Q$$
$$\frac{\partial p}{\partial \nu} = 0 \quad , \text{ on } \Sigma$$
$$p(x,T) = 0 \quad , \text{ in } \Omega$$

Remark: The first term in the expression for $J'_\beta(\lambda)$ can be formally rewritten as

$$\int_Q \left[\sum_{j,k=1}^{n} \delta a_{jk}\frac{\partial u}{\partial x_k}\frac{\partial p}{\partial x_j} + \delta a_o u p\right] dxdt$$

by using Green's formula

4. Numerical Minimization of the Smoothing Functional

(a) Gradient methods (infinite-dimensional)

The minimization of $J_\beta(\lambda)$ can be conveniently carried out by a gradient method ([2],[6]), in which J_β is iteratively minimized along the gradient direction,

[1]Lions and Magenes use sharper regularity conditions for Γ and the coefficients of A. However, the result remains unaltered. See Remark 6.1 in [11, p. 35] and Theorem 3.3 in [3, p. 32].

$\partial J_\beta/\partial\lambda$, which is defined as the unique element $\phi \in \mathscr{R}$ satisfying $J_\beta'(\lambda)\cdot h = (\phi,h)_{\mathscr{R}} \quad \forall h \in \mathscr{R}$. The calculation of ϕ reduces to an (in general) partial differential equation, which has to be numerically solved in each iteration, together with the state and adjoint equations.

This is the so-called steepest descent method. Higher order gradient methods are also applicable, but the computational effort is prohibitive.

(b) Discretized minimization methods

Let $\{\phi_i\}_{i=1}^{\infty}$ be a complete orthonormal basis for $\mathscr{R}$ and $\mathscr{R}_N$ the subspace of $\mathscr{R}$ spanned by $\phi_1,\dots,\phi_N$. $\mathscr{R}_N$ is a finite-dimensional closed subspace of $\mathscr{R}$ and $\lim_{N\to\infty} d(\lambda,\mathscr{R}_N) = 0 \quad \forall\lambda \in \mathscr{R}$

For each N, let λ_N satisfy

$$J_\beta(\lambda_N) \le J_\beta(\lambda) \qquad \forall\lambda \in \mathscr{R}_N \cap \mathscr{R}_{ad}$$

Then all weak limit points of $\{\lambda_N\}$ minimize $J_\beta(\lambda)$ over $\mathscr{R}_{ad}$. Thus, for sufficiently large N, one can solve the corresponding finite-dimensional minimization problem and obtain an approximate solution of the infinite-dimensional problem.

This is the Ritz method ([6]). A variant of this method, which uses piecewise-polynomial approximations in Sobolev spaces, seems to be computationally very attractive.

References

[1] A. BENSOUSSAN, J. L. LIONS AND G. PAPANICOLAOU, *Asymptotic Analysis for Periodic Structures*, North Holland, Amsterdam, 1978

[2] J. CÉA, *Optimization: Théorie et Algorithmes*, Dunod, Paris, 1971.

[3] G. CHAVENT, *Analyse Fonctionelle et Identification de Coefficients Répartis dans les Equations aux Dérivées Partielles*, Thèse d' Etat, Faculté des Sciences de Paris, 1971.

[4] G. CHAVENT, *About the Stability of the Optimal Control Solution of Inverse Problems*, Inverse and Improperly Posed Problems in Differential Equations, G. Anger, ed., Akademie-Verlag, Berlin, 1979, pp. 45-78.

[5] W. H. CHEN AND J. H. SEINFELD, *Estimation of Spatially Varying Parameters in Partial Differential Equations*, Int. J. Control, 15 (1972), pp. 487-495.

[6] J. W. DANIEL, *The Approximate Minimization of Functionals*, Prentice Hall, Englewood Cliffs, 1971.

[7] C. KRAVARIS, *Identification of Spatially-Varying Parameters in Distributed Parameter Systems*, Ph.D. Thesis, California Institute of Technology, Pasadena, 1984.

[8] M. M. LAVRENTIEV, V. G. ROMANOV AND V. G. VASILIEV, *Multidimensional Inverse Problems for Differential Equations*, Springer-Verlag, Berlin, 1970.

[9] J. L. LIONS, *Some Aspects of Modeling Problems in Distributed Parameter Systems*, Proceedings of the IFIP Working Conference held in Rome 1976, A. Ruberti, ed., Lecture Notes in Control and Information Sciences Vol. 1, Springer-Verlag, Berlin, 1978, pp. 11-41.

[10] J. L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.

[11] J. L. LIONS AND E. MAGENES, *Non-Homogeneous Boundary Value Problems and Applications*, Vol. II, Springer-Verlag, Berlin, 1972.

[12] M. Z. NASHED, *Approximate Regularized Solutions to Improperly Posed Linear Integral and Operator Equations*, Constructive and Computational Methods for Differential and Integral Equations, D. L. Colton and R. P. Gilbert, eds., Springer-Verlag, Berlin, 1974, pp. 289-332.

[13] A. N. TIKHONOV, *Solution of Ill-Posed Problems and the Regularization Method*, Dokl. Akad. Nauk. SSSR, 151 (1963), pp. 501-504. Soviet Math. Dokl. 4(1963), pp. 1035-1038.

[14] A. N. TIKHONOV, *Regularization of Ill-Posed Problems*, Dokl. Akad. Nauk. SSSR, 153 (1963), pp. 49-52, Soviet Math. Dokl. 4(1963), pp. 1624-1627.

[15] A. N. TIKHONOV AND V. YA. ARSENIN, *Solutions of Ill-Posed Problems*, Winston-Wiley, New York, 1977.

[16] G. WAHBA, *Practical Approximate Solutions to Linear Operator Equations When Data are Noisy*, SIAM. J. Numer. Anal., 14 (1977), pp. 651-667.